# Project description

This project includes the following folders:

1. data:
   1. fodl\_math\_test.json: tập test
   2. fodl\_math\_val.json: set val
   3. fodl\_math\_train\_imputed.json: train set with explanation data completed using GPT-3.5
   4. math\_test.json: test set similar to fodl\_math\_test.json but with an additional id field
   5. meta.pkl: file stores the data structures needed to perform tokenization
2. notebooks:
   1. random\_chance.ipynb: choose random answer
   2. fodl\_math\_exp\_0.ipynb: CNN-1D QA
   3. OpenAI\_zero\_shot.ipynb: bare LLM zeroshot learning
   4. OpenAI\_one\_shot.ipynb: bare LLM one shot learning
   5. OpenAI\_few\_shot.ipynb: bare LLM few shot learning
   6. OpenAI\_CoT\_zero\_shot.ipynb: chain-of-thoughts zero shot learning
   7. OpenAI\_CoT\_one\_shot.ipynb: chain-of-thoughts one shot learning
   8. OpenAI\_CoT\_few\_shot.ipynb: chain-of-thoughts few shot learning
3. misc:
   1. OpenAI\_generate\_explanation: missing explanation generation source code using GPT-3.5
4. predict:
   1. .dat files store data structures returned from API calls to the GPT-3.5 service
   2. The .json files save predictions as json files, which are the output of the above notebooks. Use notebook convert.ipynb to convert these files to csv format
   3. .csv files are files submitted to the Zalo AI Challenge contest
5. checkpoint:
   1. 3 checkpoint files include weight, hyperparameters and metrics when training the CNN-1D QA model

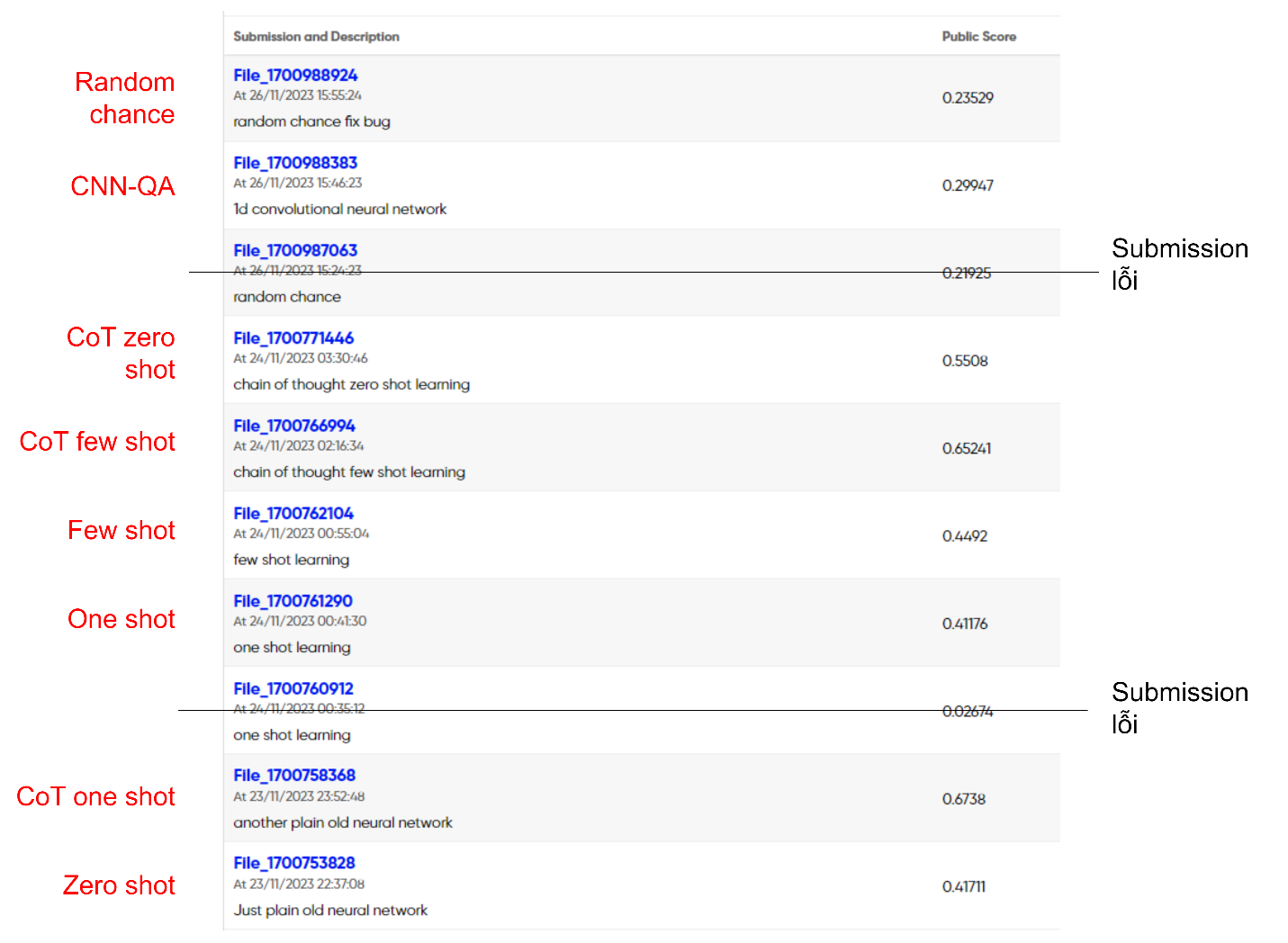
# Use notebook files

Prepare the workspace on Colab as follows using the file fodl\_math\_exp\_0.ipynb (the logs folder and .pdf files are the results after running the notebook). In case of using bare LLM and chain-of-thoughts, just upload the math\_test.json file

A screenshot of a computer

Description automatically generated

# Demonstrate test results



*Figure 1: Demonstration of the results of tests on Zalo AI Challenge 2023*

A screenshot of a computer

Description automatically generated

*Figure 2: The ranking of the best method (CoT one shot) is top 56 and the worst method (CNN-QA) is top 140 out of 160 competing teams, as of December 3, 2023*